

# Noise suppression by quantum control before and after the noise

Hiroaki Wakamura,<sup>1,\*</sup> Ryûitirô Kawakubo,<sup>1,†</sup> and Tatsuhiko Koike<sup>1,2,‡</sup>

<sup>1</sup>*Department of Physics, Keio University, Yokohama 223-8522, Japan*

<sup>2</sup>*Research and Education Center for Natural Sciences, Keio University, Yokohama 223-8521, Japan*

(Dated: January, 2017)

We discuss the possibility of protecting the state of a quantum system that goes through noise, by measurements/operations before and after the noise process. The aim is to seek for the optimal protocol that makes the input and output states as close as possible and clarify the role of the measurements therein. We consider two cases; one can perform quantum measurements/operations (i) only after the noise process and (ii) both before and after that. We prove in the two-dimensional Hilbert space that, in the case (i), the noise suppression is essentially impossible for all types of noise and, in the case (ii), the optimal protocol for the depolarizing noise is either the “do nothing” protocol or the “discriminate & reprepare” protocol. These protocols are not “truly quantum” and can be considered as classical. They involve no measurement or only use the measurement outcomes. These results describe the fundamental limitations in quantum mechanics from the viewpoint of control theory. Finally, we conjecture that a statement similar to the case (ii) holds for higher-dimensional Hilbert spaces and present some numerical evidence.

PACS numbers: 03.65.Ta, 03.67.-a, 03.67.Pp, 02.30.Yy

## I. INTRODUCTION

Measurement in quantum theory substantially differs from that in classical theory. One cannot identify the state of a system by measurement on a single sample. In addition, measurement always disturbs the system. The limitation of manipulating quantum systems can be understood as being imposed by these characteristics of quantum measurement. An example is the no-cloning theorem [1] which states that it is impossible to create identical copies of an arbitrary quantum state. If one could make a clone, then one could extract complete information from a single state by creating infinitely many copies thereof, which contradicts quantum mechanics. Other examples are the facts that one cannot discriminate non-orthogonal states perfectly and that one cannot measure non-commuting observables without errors. In presence of such impossibility, many researchers study how well one can perform these tasks mentioned above. Imperfect cloning [2, 3], state discrimination [4, 5] and uncertainty relations for noise and disturbance [6, 7] are examples of the studies that make a quantitative assessment of ability to realize the tasks approximately.

We would like to discuss an aspect of the limitations of quantum operation that one cannot protect states against noise. Here we use the word “noise” in a wide sense so that it refers to any irreversible dynamics induced by environments. In classical systems, one can protect a state against the irreversible dynamics (noise) if accurate measurements and operations can be done and if the state is not a statistical mixture, by taking the complete record of the state before the noise affects

the system. In quantum theory, it is not the case even if the state is pure. Measurement cannot be done accurately and disturbs the state. Nevertheless, one can still consider operations which approximately reverse the noise. Such approximate operations reveal the limit beyond which the noise cannot be suppressed any further. It also is interesting to understand the role which measurement plays for the task.

Noise suppression, the attempt to protect a certain class of states against given noise, is an important problem in the field of quantum control. Many researchers are working on the problem in various ways. Some try to protect a few states, while others try to protect all the pure states. The approaches are further classified by whether one uses *ex-post* control only or *ex-ante* and *ex-post* control together, where we mean by *ex-ante* and *ex-post* control the quantum measurements/operations performed before and after, respectively, the noise process. For the problem of protecting two states, Brańczyk *et al.* [8] obtained the optimal *ex-post* control for the dephasing noise. After a while, Mendonça *et al.* [9] gave a method for constructing the optimal *ex-post* control for arbitrary noise. Compared to the protection of two states, protecting all the pure states is more difficult and challenging. Zhang *et al.* [10] pointed out a part of the difficulty; they prove that *ex-post* control alone cannot suppress the depolarizing noise at all. Korotkov and Keane [11] considered *ex-ante* control as well as *ex-post* control and found that it is possible to protect, to some extent, all the pure states of a qubit against the amplitude damping noise. Along this line, Wang *et al.* [12] made a further study using numerical methods. We remark that there are still many other approaches to reducing the effect of decoherence in a wider context [13–17].

In this paper, we discuss protection of all the pure states against noise by *ex-ante* and *ex-post* control (see Fig. 1). First, we restrict ourselves to the scheme with *ex-*

\* hwakamura@rk.phys.keio.ac.jp

† rkawakub@rk.phys.keio.ac.jp

‡ koike@phys.keio.ac.jp

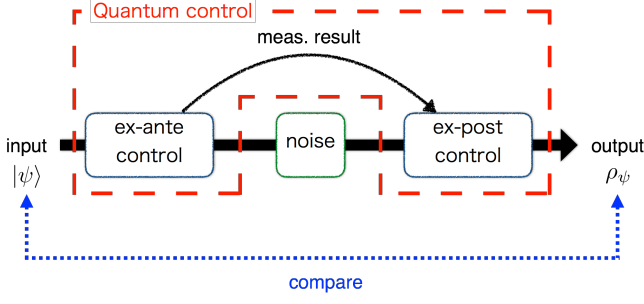


FIG. 1. The schematic diagram of the quantum control (ex-ante control and ex-post control).

post control alone and discuss if it can suppress the effect of noise. In the case of the two-dimensional Hilbert space (a qubit), we prove that the suppression is impossible for any type of noise (Theorem 1). This result, generalizing a part of Ref. [10], reveals the decisive necessity of ex-ante control if one wishes to protect all the pure states. Next, we consider the scheme with both ex-ante and ex-post control. We focus on the depolarizing noise, whose isotropic property make the suppression task more difficult, and thus suitable to see the improvement thanks to ex-ante control. Then we show, in Theorem 2, that the optimal control protocols are rather simple and easy to understand classically and ex-ante control is useful only when the noise is strong (Theorem 2). Finally, we conjecture that essentially the same results hold in the case of higher dimensional Hilbert spaces, which is supported by a numerical calculation in the case of three-dimensional Hilbert space.

The paper is organized as follows. In Sec. II, the mathematical tools used in our discussion are reviewed. In Sec. III, the setup of quantum control with ex-ante and ex-post control is explained. In Sec. IV, noise suppression for a qubit by ex-post control alone is discussed in general. We consider in Sec. V ex-ante-ex-post control for a qubit under the depolarizing noise and present a result on the optimal protocols, which is proved in Sec. VI. In Sec. VII, we propose a conjecture for higher-dimensional Hilbert spaces and show some numerical evidence. Sec. VIII is devoted to conclusion and discussions. The Appendix lists the theorems used in the text.

## II. BASICS OF QUANTUM OPERATIONS

In this section, we shall introduce basic mathematical tools and notation used in our analysis. Throughout the paper, we consider physical systems which are represented by a finite-dimensional Hilbert space. Let  $\mathcal{H}$  be such a Hilbert space. Let  $\mathcal{L}(\mathcal{H})$  be the set of all linear operators on  $\mathcal{H}$ .

An operator  $O \in \mathcal{L}(\mathcal{H})$  is said positive and denoted by  $O \geq 0$  if  $\langle \psi | O | \psi \rangle \geq 0$  holds for any  $|\psi\rangle \in \mathcal{H}$ . A quantum state is described by a density operator  $\rho \in \mathcal{L}(\mathcal{H})$  such that  $\rho \geq 0$  and  $\text{Tr } \rho = 1$ . *Fidelity* measures closeness of

two states and is defined by

$$F(\rho, |\psi\rangle\langle\psi|) := \langle \psi | \rho | \psi \rangle \quad (1)$$

in the case that at least one of the two states is pure [18]. The fidelity represents the probability of measuring  $|\psi\rangle$  when the state is  $\rho$ , and it is equal to unity if and only if  $\rho = |\psi\rangle\langle\psi|$ .

A linear map  $\mathcal{E} : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H})$  is said *positive* if  $O \geq 0$  implies  $\mathcal{E}(O) \geq 0$ . The map  $\mathcal{E}$  is said *completely positive (CP)* if the map  $\mathcal{E} \otimes \text{id}_n$  is positive for every positive integer  $n$ , where  $\text{id}_n$  denotes the identity map on  $\mathcal{L}(\mathbb{C}^n) = \mathbb{C}^{n \times n}$ . The map  $\mathcal{E}$  is said *trace-preserving* if  $\text{Tr } \mathcal{E}(O) = \text{Tr } O$  for any  $O \in \mathcal{L}(\mathcal{H})$ . A trace-preserving completely positive map is called a *TPCP map*. It is known that any physical evolution of a quantum state corresponds to a TPCP map, and vice versa [18].

An example of CP map is

$$\mathcal{A}_A(O) := AOA^\dagger \quad (2)$$

defined for  $A \in \mathcal{L}(\mathcal{H})$ . It is known that any CP map can be expressed as a sum of such  $\mathcal{A}_A$  (the Kraus representation). The CP map  $\mathcal{A}_A$  is trace-preserving only if  $A$  is a unitary operator. Another example of TPCP map, which is important in our discussion below, is mixing with the completely mixed state, or *depolarizing noise*  $\mathcal{D}_\varepsilon$ , defined by

$$\mathcal{D}_\varepsilon(\rho) := (1 - \varepsilon)\rho + \varepsilon \frac{1}{d} \text{Tr } \rho, \quad (3)$$

where  $\varepsilon$  is a parameter between 0 and 1,  $d$  is the dimensionality of  $\mathcal{H}$  and the operator  $1/d$  is the completely mixed state. The map outputs the completely mixed state with probability  $\varepsilon$  and leaves the input state untouched with probability  $1 - \varepsilon$ .

A family  $\{\mathcal{J}_\omega\}_{\omega \in \Omega}$  of CP maps with  $\sum_{\omega \in \Omega} \mathcal{J}_\omega$  being trace-preserving is called a *CP instrument*. It is known that any physical measuring process corresponds to a CP instrument, and vice versa [19]. In this paper, we assume that the number of the measurement outcomes is finite. The state evolution by the measurement is described as

$$\rho \mapsto \frac{\mathcal{J}_\omega(\rho)}{\text{Tr } \mathcal{J}_\omega(\rho)}, \quad \text{with probability } \text{Tr } \mathcal{J}_\omega(\rho). \quad (4)$$

A set  $\{M_\omega\}_{\omega \in \Omega}$  of positive operators on  $\mathcal{H}$  such that  $\sum_{\omega \in \Omega} M_\omega = 1$  is called a *POVM*. A CP instrument  $\{\text{Tr } \mathcal{J}_\omega\}_{\omega \in \Omega}$  defines a POVM by  $\text{Tr } \mathcal{J}_\omega(\rho) = \text{Tr } \rho M_\omega$ . We shall say that such a POVM  $\{M_\omega\}_{\omega \in \Omega}$  and a CP instrument  $\{\mathcal{J}_\omega\}_{\omega \in \Omega}$  are associated with each other. A POVM has the information on all statistical properties of the measurement outcomes, while a CP instrument  $\{\mathcal{J}_\omega\}$  has still more information on the state after the measurement. Any CP instrument associated with  $\{M_\omega\}_{\omega \in \Omega}$  can be written as (See Hayashi [20], p.189, Theorem 7.2)

$$\mathcal{J}_\omega = \mathcal{K}_\omega \circ \mathcal{A}_{\sqrt{M_\omega}}, \quad (5)$$

where  $\mathcal{K}_\omega$  is a TPCP map. We will call  $\{\mathcal{A}_{\sqrt{M_\omega}}\}_{\omega \in \Omega}$  a simple CP instrument associated with  $\{M_\omega\}_{\omega \in \Omega}$ .

The set  $\mathcal{L}(\mathcal{H})$  can be regarded as a Hilbert space with the Hilbert-Schmidt inner product  $\langle X, Y \rangle_{\text{HS}} := \text{Tr } X^\dagger Y$ . Then a linear map  $\mathcal{E}$  on  $\mathcal{L}(\mathcal{H})$  is a linear operator on the Hilbert space  $\mathcal{L}(\mathcal{H})$ . Thus the trace of  $\mathcal{E} \in \mathcal{L}(\mathcal{L}(\mathcal{H}))$  is defined as

$$\text{Tr}_{\text{HS}} \mathcal{E} := \sum_i \langle V_i, \mathcal{E}(V_i) \rangle_{\text{HS}}, \quad (6)$$

where  $\{V_i\}_i$  is an orthonormal basis of the Hilbert space  $\mathcal{L}(\mathcal{H})$ . For example, when  $\dim \mathcal{H} = 2$ , the set of Pauli operators  $\{\sigma_\mu / \sqrt{2}\}_{\mu=0}^3$ , with  $\sigma_0$  being the identity operator, is an orthonormal basis of  $\mathcal{L}(\mathcal{H})$  so that the trace of  $\mathcal{E} \in \mathcal{L}(\mathcal{L}(\mathcal{H}))$  can be written as

$$\text{Tr}_{\text{HS}} \mathcal{E} = \frac{1}{2} \sum_{\mu=0}^3 \text{Tr} [\sigma_\mu \mathcal{E}(\sigma_\mu)]. \quad (7)$$

### III. THE SETUP

In this section, we shall present the main problem and its mathematical formulation.

We consider the *ex-ante-ex-post quantum control scheme* defined by the following sequence of processes (depicted in Fig. 1):

1. “state preparation”  
An unknown state  $|\psi\rangle \in \mathcal{H}$  is prepared.
2. “ex-ante control”  
A measurement is performed, which is described by a CP instrument  $\{\mathcal{J}_\omega\}_{\omega \in \{1, \dots, M\}}$ , where  $M$  is a positive integer.
3. “noise”  
The state undergoes an undesired evolution, called “noise,” described by a TPCP map  $\mathcal{N}$ .
4. “ex-post control”  
An operation, which depends on the measurement outcome  $\omega$  of the ex-ante control, is performed on the system. This is described by a family of TPCP maps,  $\{\mathcal{C}_\omega\}_{\omega \in \{1, \dots, M\}}$ , which we call the ex-post control.

For given noise  $\mathcal{N}$ , an ex-ante-ex-post control protocol is specified by the family  $\{(\mathcal{J}_\omega, \mathcal{C}_\omega)\}_{\omega \in \{1, \dots, M\}}$ . We assume throughout the paper that the prepared state is pure, though one can consider more general mixed state preparation. We also assume that the state  $|\psi\rangle$  above is *completely unknown*, i.e., the probability distribution is uniform on the unit sphere in  $\mathcal{H}$ .

The problem that we want to consider is, for given noise  $\mathcal{N}$ , to find an *optimal* scheme such that the states after the measurement with outcome  $\omega$  are as similar to the original state  $|\psi\rangle\langle\psi|$  as possible. For defining the optimality, it is natural to introduce some evaluation function  $h$  and take an average with respect to the probability of obtaining  $\omega$ , and then take an average with respect to  $|\psi\rangle$  which is completely unknown,

$$\begin{aligned} & \int_{\| |\psi\rangle \| = 1} d\psi \sum_{\text{measurement outcomes}} \text{evaluation function}(\text{final state}, \text{initial state}) \times \text{probability} \\ &= \int_{\| |\psi\rangle \| = 1} d\psi \sum_{\omega} h \left( \mathcal{C}_\omega \circ \mathcal{N} \left( \frac{\mathcal{J}_\omega(|\psi\rangle\langle\psi|)}{\text{Tr } \mathcal{J}_\omega(|\psi\rangle\langle\psi|)} \right), |\psi\rangle\langle\psi| \right) \text{Tr } \mathcal{J}_\omega(|\psi\rangle\langle\psi|), \end{aligned} \quad (8)$$

where the integral is over all unit vectors in  $\mathcal{H}$  with the uniform measure normalized by  $\int_{\| |\psi\rangle \| = 1} d\psi = 1$ . We choose the function  $h$  to be the fidelity  $F$  in Eq. (1). An advantage of the choice is that the resulting total evaluation function, the *average fidelity*

$$\bar{F} = \int_{\| |\psi\rangle \| = 1} d\psi \langle \psi | \mathcal{E}(|\psi\rangle\langle\psi|) | \psi \rangle, \quad (9)$$

depends on the protocol  $\{(\mathcal{J}_\omega, \mathcal{C}_\omega)\}_\omega$  only through the *average operation*

$$\mathcal{E} := \sum_{\omega=1}^M \mathcal{C}_\omega \circ \mathcal{N} \circ \mathcal{J}_\omega \quad (10)$$

which is a TPCP map.

We close the section by presenting a useful formula for the average fidelity.

**Lemma 1.** *The average fidelity  $\bar{F}$  is given by*

$$\bar{F} = \frac{1}{d(d+1)} (d + \text{Tr}_{\text{HS}} \mathcal{E}), \quad (11)$$

where  $\mathcal{E}$  is the average operation (10) of the protocol  $\{(\mathcal{J}_\omega, \mathcal{C}_\omega)\}_\omega$  and  $d := \dim \mathcal{H}$ .

*Proof.* Let  $S$  be the swap operator,  $S(|\psi\rangle|\phi\rangle) := |\phi\rangle|\psi\rangle$ , or in an orthonormal basis  $\{|ij\rangle\}_{i,j}$ ,  $S = \sum_{ij} |ij\rangle\langle ji|$ .

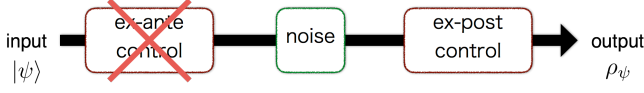


FIG. 2. The schematic diagram of the quantum control (only ex-post control).

One can easily see that

$$\text{Tr}[AB] = \text{Tr}[S(A \otimes B)]. \quad (12)$$

It follows from (9) and (12) that

$$\bar{F} = \text{Tr}[S(\text{id} \otimes \mathcal{E})(Q)], \quad (13)$$

where  $Q := \int_{\|\psi\|=1} d\psi |\psi\rangle\langle\psi| \otimes |\psi\rangle\langle\psi|$ . Then the result is easily obtained by the following formulas,

$$Q = \frac{1+S}{d(d+1)}, \quad (14)$$

$$\text{Tr}[S(\text{id} \otimes \mathcal{E})(S)] = \text{Tr}_{\text{HS}} \mathcal{E}. \quad (15)$$

Indeed, from (12), (13), (14) and (15), one has

$$\begin{aligned} (d(d+1))\bar{F} &= \text{Tr}[S(\text{id} \otimes \mathcal{E})(1)] + \text{Tr}[S(\text{id} \otimes \mathcal{E})(S)] \\ &= \text{Tr} \mathcal{E}(1) + \text{Tr}_{\text{HS}} \mathcal{E} = d + \text{Tr}_{\text{HS}} \mathcal{E}. \end{aligned} \quad (16)$$

Let us show (14) and (15). Eq. (14) is seen [21] by noting that  $Q$  commutes with  $U \otimes U$  for any  $U$  in  $\text{SU}(d)$ , the special unitary group on  $\mathbb{C}^d$ . By Schur's lemma,  $Q$  acts as scalar operators on the symmetric and antisymmetric subspaces of  $\mathcal{H} \otimes \mathcal{H}$ , which are the spaces of irreducible representations. The scalar factors are found easily. Eq. (15) is seen by direct calculation,  $\text{LHS} = \sum_{ijkl} \langle ij | (|k\rangle\langle l| \otimes \mathcal{E}(|l\rangle\langle k|)) | ji \rangle = \sum_{ij} \langle j | \mathcal{E}(|j\rangle\langle i|) | i \rangle = \text{RHS}$ . This completes the proof.  $\square$

#### IV. EX-POST CONTROL

In this section, we shall consider the noise suppression by ex-post control only (Fig. 2) and present our first main result.

The control sequence has no branches and the protocol is determined by a TPCP map  $\mathcal{C}$  describing the ex-post control. Thus our aim is to find the optimal ex-post control  $\mathcal{C}$ . The average fidelity to be maximized is

$$\bar{F} = \frac{1}{d(d+1)} (d + \text{Tr}_{\text{HS}} \mathcal{C} \circ \mathcal{N}), \quad (17)$$

from Lemma 1.

**Theorem 1.** *For any noise  $\mathcal{N}$  in the two-dimensional Hilbert space  $\mathcal{H}$ , the optimal ex-post control protocol is a unitary transformation.*

*Proof.* In the two-dimensional Hilbert space  $\mathcal{H}$ , we can express a general TPCP map  $\mathcal{E}$  in the basis  $\{\sigma_\mu\}_{0 \leq \mu \leq 3}$

such that

$$\mathcal{E}(1) = 1 + \sum_{k=1}^3 t_k \sigma_k, \quad \mathcal{E}(\sigma_j) = \sum_{j,k=1}^3 s_{jk} \sigma_k. \quad (18)$$

A necessary condition for positivity and trace-preserving property is  $\mathcal{E}(1 \pm \sigma_j) \geq 0$ , or

$$\sum_{k=1}^3 |t_k \pm s_{jk}|^2 \leq 1. \quad (19)$$

In the two-dimensional Hilbert space, one can transform the noise  $\mathcal{N}$  to  $\tilde{\mathcal{N}}$  by unitary operations for the input and output states such that  $\tilde{\mathcal{N}}$  has diagonal  $s_{jk}$ , namely, one has

$$\mathcal{N} = \mathcal{A}_U \circ \tilde{\mathcal{N}} \circ \mathcal{A}_V \quad (20)$$

with some unitary operators  $U$  and  $V$  on  $\mathcal{H}$ , where  $\mathcal{A}$  is defined in (2). Introducing  $\tilde{\mathcal{C}} := \mathcal{A}_V \circ \mathcal{C} \circ \mathcal{A}_U$ , one has

$$\begin{aligned} \text{Tr}_{\text{HS}} \mathcal{C} \circ \mathcal{N} &= \text{Tr}_{\text{HS}} \tilde{\mathcal{C}} \circ \tilde{\mathcal{N}} = \frac{1}{2} \sum_{\mu=0}^3 \text{Tr} \sigma_\mu \tilde{\mathcal{C}} \circ \tilde{\mathcal{N}}(\sigma_\mu) \\ &= 1 + \frac{1}{4} \sum_{j=1}^3 \text{Tr} \sigma_j \tilde{\mathcal{N}}(\sigma_j) \text{Tr} \sigma_j \tilde{\mathcal{C}}(\sigma_j), \end{aligned} \quad (21)$$

where we have used (7) and trace preserving property of  $\tilde{\mathcal{C}}$ . Because  $\tilde{\mathcal{C}}$  is TPCP, (19) implies  $\text{Tr} \sigma_j \tilde{\mathcal{C}}(\sigma_j) \leq 2$  for  $j = 1, 2, 3$ , so that

$$\text{Tr}_{\text{HS}} \mathcal{C} \circ \mathcal{N} \leq 1 + \frac{1}{2} \sum_{j=1}^3 \text{Tr} \sigma_j \tilde{\mathcal{N}}(\sigma_j). \quad (22)$$

The equality holds only when  $\text{Tr} \sigma_j \tilde{\mathcal{C}}(\sigma_j) = 2$ , so that, by (19) again, one has  $\tilde{\mathcal{C}} = \text{id}$ . Thus one has  $\mathcal{C} = \mathcal{A}_V^{-1} \circ \mathcal{A}_U^{-1}$ , which is a unitary transformation. Then the maximum average fidelity  $\bar{F}$  is  $(6 + \sum_j \text{Tr} \sigma_j \tilde{\mathcal{N}}(\sigma_j))/12$ .  $\square$

Intuitively, this result says that essentially no ex-post control can suppress the effect of noise if we are completely ignorant of the initial state. On the other hand, when we have some knowledge of the initial state then we can suppress the effect of noise by quantum control [8, 9]. Instead of restricting the candidates for initial state, we consider ex-ante control to extract some information of the initial state, which is discussed in the next section.

From the proof, we find that the quantity  $(1/2) \sum_j \text{Tr} \sigma_j \tilde{\mathcal{N}}(\sigma_j)$  characterizes to what extent the noise is reversible and allows a geometrical interpretation. In the Bloch sphere representation of the state, the whole sphere is mapped to an ellipsoid. Each  $(1/2) \text{Tr} \sigma_j \tilde{\mathcal{N}}(\sigma_j)$  is the contraction rate along a principal axis.

In closing this section, we remark that an ex-ante-ex-post control with a single branch is equivalent to an ex-post control. In fact, the average fidelity can be written

as

$$\bar{F} = \frac{1}{d(d+1)} \left( d + \text{Tr}_{\text{HS}} \mathcal{J} \circ \mathcal{C} \circ \mathcal{N} \right) \quad (23)$$

by Lemma 1 and the cyclic property of the trace. Here,  $\mathcal{J}, \mathcal{C}$  are TPCP maps corresponding to ex-ante control and ex-post control, respectively, but the same fidelity can be achieved by an ex-post control  $\mathcal{J} \circ \mathcal{C}$ .

## V. EX-ANTE-EX-POST CONTROL

In this section, we discuss the noise suppression by ex-ante-ex-post control (Fig. 1) on a qubit. We shall consider the case of two branches because a POVM consisting of two projections gives the optimal discrimination between completely unknown states of a qubit [22].

In general, there is a trade-off between the information gained and the disturbance caused by the ex-ante control. The ex-ante control can extract some information on the initial state which may be useful for noise suppression. At the same time, the measurement disturbs the state. Thus one might expect that a protocol with a “soft” ex-ante measurement could be optimal. It turns out, however, that this is not the case.

**Theorem 2.** *Let the noise  $\mathcal{N}$  be the depolarizing noise  $\mathcal{D}_\varepsilon$  defined in (3). For the two-dimensional Hilbert space  $\mathcal{H}$ , the optimal ex-ante-ex-post control protocol  $\{(\mathcal{J}_\omega, \mathcal{C}_\omega)\}_{\omega=1,2}$  is given as follows.*

(i) *When the noise is weak,  $\varepsilon \leq 2/3$ , the “do nothing” protocol is optimal, which is given by*

$$\mathcal{J}_\omega \propto \text{id}, \quad \mathcal{C}_\omega = \text{id}, \quad \omega = 1, 2. \quad (24)$$

*The optimal average fidelity is  $\bar{F}_{\text{DN}} = 1 - \varepsilon/2$ .*

(ii) *When the noise is strong,  $\varepsilon \geq 2/3$ , the “discriminate & reprepare” protocol is optimal, which is given by*

$$\mathcal{J}_\omega(\rho) = |\phi_\omega\rangle\langle\phi_\omega| \rho |\phi_\omega\rangle\langle\phi_\omega|, \quad (25)$$

$$\mathcal{C}_\omega(\rho) = |\phi_\omega\rangle\langle\phi_\omega| \text{Tr} \rho, \quad (26)$$

where  $\{|\phi_\omega\rangle\}_{\omega=1,2}$  is an arbitrary orthonormal basis of  $\mathcal{H}$ . The optimal average fidelity is  $\bar{F}_{\text{DR}} = 2/3$ .

We give the proof in the next section.

The “do nothing” protocol literally does nothing, and merely lets the system undergo the noise. The value of the average fidelity  $\bar{F}_{\text{DN}}$  in Theorem 2 can be obtained by direct calculation. Namely, one substitutes  $\mathcal{E} = \mathcal{D}_\varepsilon$  into the definition (9) of  $\bar{F}$  and has

$$\bar{F}_{\text{DN}} = \int_{\|\psi\|=1} \left( (1-\varepsilon) \langle\psi|\psi\rangle\langle\psi|\psi\rangle + \varepsilon \frac{1}{2} \langle\psi|\psi\rangle \right) = 1 - \frac{\varepsilon}{2}. \quad (27)$$

The “discriminate & reprepare” protocol means that one measures and discriminates between a certain two orthogonal states before the noise process and reprepares

the corresponding state after the noise process. The value of the average fidelity  $\bar{F}_{\text{DR}}$  in Theorem 2 can be calculated as follows. Without loss of generality, one can choose  $|\phi_\omega\rangle = |\omega\rangle$ ,  $\omega = 0, 1$ , where  $\sigma_3|0\rangle = |0\rangle$  and  $\sigma_3|1\rangle = -|1\rangle$ . Substituting  $\langle\omega|\sigma_3|\omega\rangle = (-1)^\omega$  and  $\langle\omega|\sigma_1|\omega\rangle = \langle\omega|\sigma_2|\omega\rangle = 0$  into (7), one obtains  $\text{Tr}_{\text{HS}} \mathcal{E} = 2$  and hence

$$\bar{F}_{\text{DR}} = \frac{1}{6} \left( 2 + \text{Tr}_{\text{HS}} \mathcal{E} \right) = \frac{2}{3} \quad (28)$$

from Lemma 1. It is known from the study of imperfect cloning [22, 23] that any “discriminate & reprepare” protocol with arbitrary number of branches does not give a larger value.

If the noise is infinitesimally weak, it is natural that doing nothing is better than the other protocols. If the noise is so strong that the state is completely destroyed after the noise process, one should obtain information on the initial state as much as possible before the system goes through the noise process. The result implies that there is no intermediate regime where the optimal protocol involves weak measurements. The “discriminate & reprepare” protocol only uses the classical information extracted by the ex-ante measurement while the “do nothing” protocol perform no quantum measurement or operation. These protocols are rather classically motivated and are not “truly quantum,” in that they do not reflect any trade-off relation between information gain and disturbance. It is remarkable that these classical control protocols are better than any other quantum control protocols. The result shows that we cannot suppress the noise even if we can perform ex-ante control. This may be understood as fundamental limitations in quantum mechanics.

## VI. PROOF OF THEOREM 2

In this section, we give a proof of Theorem 2, after showing two lemmas.

**Lemma 2.** *Let  $\mathcal{H}$  be two-dimensional. For any TPCP map  $\mathcal{E}$ , there exists a TPCP map  $\tilde{\mathcal{E}}$  that satisfies*

$$\mathcal{D}_\varepsilon \circ \mathcal{E} = \tilde{\mathcal{E}} \circ \mathcal{D}_\varepsilon. \quad (29)$$

*Proof.* The cases  $\varepsilon = 0, 1$  are trivial so that we assume  $\varepsilon \neq 0, 1$ . Then there exists  $\mathcal{D}_\varepsilon^{-1}$  which is actually  $\mathcal{D}_{-\varepsilon/(1-\varepsilon)}$ . One therefore has

$$\tilde{\mathcal{E}} = \mathcal{D}_\varepsilon \circ \mathcal{E} \circ \mathcal{D}_{-\varepsilon/(1-\varepsilon)}. \quad (30)$$

We prove that the map  $\tilde{\mathcal{E}}$  is CP, though  $\mathcal{D}_{-\varepsilon/(1-\varepsilon)}$  is not. Any Hermiticity-preserving linear map  $\sum_{\mu=0}^3 x^\mu \sigma_\mu \mapsto \sum_{\mu=0}^3 x'^\mu \sigma_\mu$  is specified by a linear map from  $\mathbb{R}^4$  to  $\mathbb{R}^4$ ,  $(x^\mu) \mapsto (x'^\mu)$ ,  $\mu = 0, 1, 2, 3$ . From a theorem by Ruskai

*et al.* (see the Appendix), the TPCP map  $\mathcal{E}$  is expressed as

$$\begin{pmatrix} x'^0 \\ \mathbf{x}' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \mathbf{t} & E \end{pmatrix} \begin{pmatrix} x^0 \\ \mathbf{x} \end{pmatrix}, \quad (31)$$

where  $\mathbf{x} = (x^1, x^2, x^3)^T$ ,  $\mathbf{t} \in \mathbb{R}^3$ ,  $E$  is a  $3 \times 3$  matrix, and  $\mathbf{t}$  and the (signed) singular values  $d_i$  of  $E$  satisfy (A.4), (A.6), (A.7) and (A.8). In the same way,  $\tilde{\mathcal{E}}$  is expressed as

$$\begin{pmatrix} x'^0 \\ \mathbf{x}' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ (1-\varepsilon)\mathbf{t} & E \end{pmatrix} \begin{pmatrix} x^0 \\ \mathbf{x} \end{pmatrix}. \quad (32)$$

Since  $0 < \varepsilon < 1$ , the components of the matrix above also satisfy the conditions (A.4), (A.6), (A.7) and (A.8). Thus  $\tilde{\mathcal{E}}$  is TPCP.  $\square$

**Lemma 3.** *Let  $A$ ,  $\tilde{R}$ ,  $D$  and  $R$  be real square matrices. If  $A$  and  $D$  are diagonal and  $\tilde{R}$  is orthogonal, the following inequality holds,*

$$\text{Tr} [ARD\tilde{R}] \leq \sum_j |d_j| \left( \sum_i a_i^2 r_{ij}^2 \right)^{1/2}, \quad (33)$$

where  $a_j$  and  $d_j$  are the diagonal elements of  $A$  and  $D$ , respectively, and  $r_{ij}$  is the  $(i, j)$  element of  $R$ .

*Proof.* Let  $\tilde{r}_{ij}$  be the  $(i, j)$  element of  $\tilde{R}$ . One has

$$\text{Tr} [ARD\tilde{R}] = \sum_{ij} a_i r_{ij} d_j \tilde{r}_{ji} = \sum_j |d_j| \sum_i a_i r_{ij} \frac{d_j \tilde{r}_{ji}}{|d_j|}. \quad (34)$$

One can view the  $i$ -sum as an inner product of vectors  $(a_i r_{ij})_i$  and  $(d_j \tilde{r}_{ji}/|d_j|)_i$ . Applying the Cauchy-Schwarz inequality to the inner product, and using the orthogonality of  $\tilde{R}$ , one is able to show the claim.  $\square$

*Proof of Theorem 2.* The proof consists of five parts. The main idea is to show that  $\tilde{F}$  does not exceed the values attained by the “do nothing” and “discriminate & repare” protocols.

*Step 1. Reduction to simple CP instrument.* By Lemma 1, the optimal protocol  $\{(\mathcal{J}_\omega, \mathcal{C}_\omega)\}_{\omega=1,2}$  is the maximizer of  $\sum_{\omega=1}^2 f_\omega$ , with

$$f_\omega := \text{Tr}_{\text{HS}} \mathcal{C}_\omega \circ \mathcal{D}_\varepsilon \circ \mathcal{J}_\omega. \quad (35)$$

As was explained in (5), the CP instrument  $\{\mathcal{J}_\omega\}$  is specified by a family  $\{\mathcal{K}_\omega\}$  of TPCP maps and a POVM  $\{M_\omega\}$  so that  $\mathcal{J}_\omega = \mathcal{K}_\omega \circ \mathcal{A}_{\sqrt{M_\omega}}$ . By Lemma 2, there is a TPCP map  $\tilde{\mathcal{K}}_\omega$  which satisfies  $\mathcal{D}_\varepsilon \circ \mathcal{K}_\omega = \tilde{\mathcal{K}}_\omega \circ \mathcal{D}_\varepsilon$ . From these,  $f_\omega$  can be written as

$$\begin{aligned} f_\omega &= \text{Tr}_{\text{HS}} \mathcal{C}_\omega \circ \mathcal{D}_\varepsilon \circ \mathcal{K}_\omega \circ \mathcal{A}_{\sqrt{M_\omega}} \\ &= \text{Tr}_{\text{HS}} \mathcal{C}_\omega \circ \tilde{\mathcal{K}}_\omega \circ \mathcal{D}_\varepsilon \circ \mathcal{A}_{\sqrt{M_\omega}} \\ &= \text{Tr}_{\text{HS}} \tilde{\mathcal{C}}_\omega \circ \mathcal{D}_\varepsilon \circ \mathcal{A}_{\sqrt{M_\omega}}, \end{aligned} \quad (36)$$

where  $\tilde{\mathcal{C}}_\omega := \mathcal{C}_\omega \circ \tilde{\mathcal{K}}_\omega$ . Thus, the original optimization problem for the protocol  $\{(\mathcal{C}_\omega, \mathcal{J}_\omega)\}$  is translated to that for a protocol  $\{(\tilde{\mathcal{C}}_\omega, \mathcal{A}_{\sqrt{M_\omega}})\}$  specified by a family of TPCP maps  $\{\tilde{\mathcal{C}}_\omega\}$  and a POVM  $\{M_\omega\}$ .

Let us choose a basis  $\{|0\rangle, |1\rangle\}$  in which  $M_1$  and  $M_2 = 1 - M_1$  are diagonal, so that one has

$$M_\omega = \alpha_\omega |0\rangle\langle 0| + \beta_\omega |1\rangle\langle 1|, \quad (37)$$

$$\alpha_\omega, \beta_\omega \geq 0, \quad \sum_\omega \alpha_\omega = \sum_\omega \beta_\omega = 1. \quad (38)$$

From (7) and (35), one has  $f_\omega = \frac{1}{2}(\text{Tr} M_\omega + I_\omega)$  so that

$$f_1 + f_2 = 1 + \frac{1}{2}(I_1 + I_2), \quad (39)$$

where

$$\begin{aligned} I_\omega &:= \sum_{j=1}^3 \text{Tr} [\sigma_j \tilde{\mathcal{C}}_\omega \circ \mathcal{D}_\varepsilon \circ \mathcal{A}_{\sqrt{M_\omega}}(\sigma_j)] \\ &= (1-\varepsilon) \sqrt{\alpha_\omega \beta_\omega} \sum_{j=1}^2 \text{Tr} [\sigma_j \tilde{\mathcal{C}}_\omega(\sigma_j)] \\ &\quad + \frac{\alpha_\omega - \beta_\omega}{2} \text{Tr} [\sigma_3 \tilde{\mathcal{C}}_\omega(1)] + (1-\varepsilon) \frac{\alpha_\omega + \beta_\omega}{2} \text{Tr} [\sigma_3 \tilde{\mathcal{C}}_\omega(\sigma_3)]. \end{aligned} \quad (40)$$

*Step 2. Necessity for each  $\tilde{\mathcal{C}}_\omega$  to be extreme.* From here to the end of Step 4, we fix the POVM  $\{M_\omega\}$  and vary the TPCP maps  $\tilde{\mathcal{C}}_\omega$  to obtain a bound for each  $I_\omega$ . At this stage one can treat each  $I_\omega$  independently. We drop the subscript  $\omega$  from the variables and parameters till Step 4.

Since  $f$  is a linear functional of  $\tilde{\mathcal{C}}$ , we observe that the optimal  $\tilde{\mathcal{C}}$  must be one of the extreme points in the convex space of TPCP maps. From a theorem by Ruskai *et al.* (see the Appendix), such a TPCP map can be written in the form (A.5) with the condition (A.9). Thus one has

$$I = (\alpha - \beta) d_0 r_{33} + (1-\varepsilon) \text{Tr} [ARD\tilde{R}], \quad (41)$$

where  $R$  and  $\tilde{R}$  are real  $3 \times 3$  rotation matrices,  $r_{33}$  is the  $(3, 3)$  element of  $R$ ,  $A := \text{diag}[2\sqrt{\alpha\beta}, 2\sqrt{\alpha\beta}, \alpha + \beta]$ ,  $D := \text{diag}[d_1, d_2, d_1 d_2]$ ,  $d_0^2 := (1 - d_1^2)(1 - d_2^2)$  and  $d_1, d_2 \in [-1, 1]$ .

*Step 3. An  $\tilde{R}$ -independent upper bound of  $I$ .* We shall derive an upper bound  $J$  of  $I$ , which is independent from  $\tilde{R}$ . From Lemma 3, one has

$$\text{Tr} [ARD\tilde{R}] \leq \sum_j |d_j| \left( \sum_i a_i^2 r_{ij}^2 \right)^{1/2}, \quad (42)$$

where  $a_i$  and  $d_i$  are diagonal elements of  $A$  and  $D$ , respectively, and  $r_{ij}$  is the  $(i, j)$  component of  $R$ .

Next, it follows from orthogonality of  $R$  that the matrix with  $(i, j)$  element being  $r_{ij}^2$  is doubly stochastic.

From the Birkhoff-von Neumann theorem (see the Appendix), such a matrix must be a convex combination of permutation matrices. Furthermore, because we are considering the case  $a_1 = a_2 (= 2\sqrt{\alpha\beta})$  in the maximization of  $I$ , it is enough to consider the case

$$(r_{ij}^2)_{0 \leq i,j \leq 3} = (1-2p) \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} + (p+q) \begin{pmatrix} & 1 & & \\ & & 1 & \\ 1 & & & \\ & & & 1 \end{pmatrix} + (p-q) \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}, \quad (43)$$

where  $0 \leq q \leq p \leq 1/2$ . From (41), (42) and (43), one has

$$\begin{aligned} I \leq J := & |(\alpha - \beta)d_0| \sqrt{1-2p} \\ & + (1-\varepsilon) \left[ |d_1| \sqrt{a_1^2 + (a_3^2 - a_1^2)(p+q)} \right. \\ & \left. + |d_2| \sqrt{a_1^2 + (a_3^2 - a_1^2)(p-q)} + |d_1 d_2| \sqrt{a_3^2 - 2(a_3^2 - a_1^2)p} \right]. \end{aligned} \quad (44)$$

The bound  $J$  is a function of  $(p, q, d_1, d_2)$  while  $a_1$  and  $a_3$  are parameters.

*Step 4. Joint concavity and a  $\tilde{\mathcal{C}}$ -independent bound.* Let  $x := (d_1^2 + d_2^2)/2$ ,  $y := (d_1^2 - d_2^2)/2$ . Then  $J(p, q, x, y)$  is jointly concave with respect to variables  $q$  and  $y$ , which can be easily seen by direct calculation of the Hessian of each term of  $J$ . Furthermore,  $J$  is invariant under  $(q, y) \mapsto (-q, -y)$ . Thus one has

$$\begin{aligned} J(p, q, x, y) &= \frac{1}{2} (J(p, q, x, y) + J(p, -q, x, -y)) \quad (45) \\ &\leq J(p, 0, x, 0) \\ &= |\alpha - \beta|(1-x)\sqrt{1-2p} \\ &\quad + (1-\varepsilon) \left[ 2\sqrt{x}\sqrt{4\alpha\beta + (\alpha - \beta)^2 p} \right. \\ &\quad \left. + x\sqrt{(\alpha + \beta)^2 - 2(\alpha - \beta)^2 p} \right], \end{aligned} \quad (46)$$

where we have substituted the expressions of  $a_1, a_2$  and  $d_0$ . Because  $\alpha, \beta \in [0, 1]$  and  $p \in [0, 1/2]$ , one can show by a simple observation that the bound (46) does not exceed

$$(1-x)\sqrt{r} + 2(1-\varepsilon)\sqrt{x}\sqrt{(\alpha + \beta)^2 - r} + (1-\varepsilon)(\alpha + \beta)x, \quad (47)$$

where  $r := (\alpha - \beta)^2(1-p) \in [0, 1]$ . Applying the Cauchy-Schwarz inequality to the first two terms of (47), one sees that (47) does not exceed

$$(\alpha + \beta) \left( \sqrt{(1-x)^2 + 4(1-\varepsilon)^2 x} + (1-\varepsilon)x \right). \quad (48)$$

This is a convex function of  $x$  hence reaches the maximum at the boundary  $x = 0, 1$ . Thus, from (44), we arrive at a  $\tilde{\mathcal{C}}$ -invariant upper bound,

$$I \leq (\alpha + \beta) \max \{3(1-\varepsilon), 1\}. \quad (49)$$

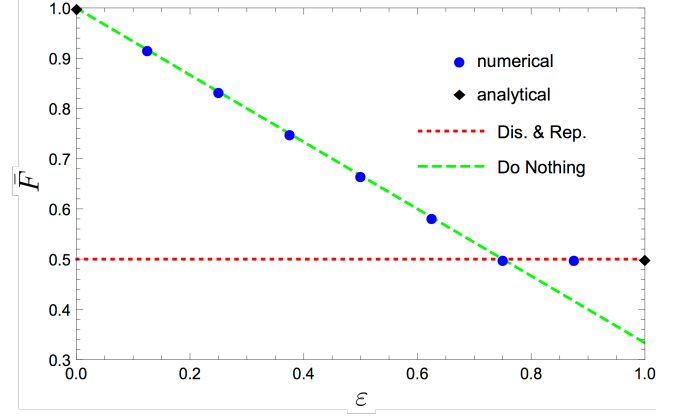


FIG. 3. (Color online) The maximum average fidelity  $\bar{F}$  achieved by ex-ante-ex-post control for the depolarizing noise of strength  $\varepsilon$ , when  $\dim \mathcal{H} = 3$ . The (blue) round points are the maxima obtained by numerical optimization, the (black) diamond points are exact values for  $\varepsilon = 0, 1$ , the (green) dashed line is  $\bar{F}$  achieved by the “do nothing” protocol, and the (red) dotted line is that obtained by the “discriminate & reprepare” protocol. The results support Conjecture 1.

*Step 5. A protocol-independent bound for  $\bar{F}$  and its attainability.* We revive the subscript  $\omega$ . Because  $\sum_{\omega} \alpha_{\omega} = \sum_{\omega} \beta_{\omega} = 1$ , one immediately obtains a protocol-independent upper bound from (49),  $I_1 + I_2 \leq 2 \max \{3(1-\varepsilon), 1\}$ . This is equivalent to

$$\bar{F} \leq \frac{1}{6} (3 + \max \{3(1-\varepsilon), 1\}). \quad (50)$$

The values  $\bar{F} = 1 - \varepsilon/2$  and  $\bar{F} = 2/3$  are attained by “do nothing” and “discriminate & reprepare” protocols, respectively, as was shown in (27) and (28).  $\square$

## VII. HIGHER DIMENSIONS

Though Theorem 2 is valid only for the two-dimensional Hilbert space  $\mathcal{H}$ , similar results may hold in higher dimensions. In this section, we propose a conjecture in general dimensionality and show a numerical evidence for the three-dimensional Hilbert space.

**Conjecture 1.** *Let  $\mathcal{D}_{\varepsilon}$  be the depolarizing noise (3) in the  $d$ -dimensional Hilbert space  $\mathcal{H}$ . Then the optimal ex-ante-ex-post control  $\{(\mathcal{I}_{\omega}, \mathcal{C}_{\omega})\}_{\omega=1}^d$  for  $\mathcal{D}_{\varepsilon}$  is given as follows.*

(i) *When the noise is weak,  $\varepsilon \leq d/(d+1)$ , the “do nothing” protocol is optimal, which is given by*

$$\mathcal{I}_{\omega} \propto \text{id}, \quad \mathcal{C}_{\omega} = \text{id}, \quad \omega = 1, \dots, d. \quad (51)$$

*The optimal average fidelity is  $\bar{F}_{DN} = 1 - (d-1)\varepsilon/d$ .*

(ii) *When the noise is strong,  $\varepsilon \geq d/(d+1)$ , the “discriminate & reprepare” protocol is optimal, which is given*

by

$$\mathcal{J}_\omega(\rho) = |\phi_\omega\rangle\langle\phi_\omega| \rho |\phi_\omega\rangle\langle\phi_\omega|, \quad (52)$$

$$\mathcal{C}_\omega(\rho) = |\phi_\omega\rangle\langle\phi_\omega| \text{Tr} \rho, \quad \omega = 1, \dots, d, \quad (53)$$

where  $\{|\phi_\omega\rangle\}_{\omega=1}^d$  is an arbitrary orthonormal basis of  $\mathcal{H}$ . The optimal average fidelity is  $\bar{F}_{DR} = 2/(d+1)$ .

The proof of Theorem 2 does not work in the same way because we used a concrete characterization of the extreme points of TPCP maps when  $\dim \mathcal{H} = 2$ .

In our numerical calculation below, we make use of Choi's correspondence (see the Appendix) which relates the CP maps  $\mathcal{C}_\omega, \mathcal{J}_\omega \in \mathcal{L}(\mathcal{L}(\mathcal{H}))$  with positive operators  $C_\omega, I_\omega \in \mathcal{L}(\mathcal{H} \otimes \mathcal{H})$ , called the Choi operators, as in (A.2).

**Lemma 4.** *The average fidelity for an ex-ante-ex-post control protocol  $\{(\mathcal{J}_\omega, \mathcal{C}_\omega)\}_{\omega=1}^d$  is given by*

$$\bar{F} = \frac{1}{d(d+1)} \left( d + \sum_{\omega=1}^d \text{Tr} [R_\omega(\text{id} \otimes \mathcal{N})(S)] \right), \quad (54)$$

where

$$R_\omega := \text{Tr}_C [(1_A \otimes C_{\omega, CB}^T)(I_{\omega, AC} \otimes 1_B)], \quad \omega = 1, \dots, d, \quad (55)$$

is the partial transpose of the Choi operator of  $\mathcal{J}_\omega \circ \mathcal{C}_\omega$  [ $\mathcal{H}_A, \mathcal{H}_B$  and  $\mathcal{H}_C$  are copies of  $\mathcal{H}$ ,  $\text{Tr}_C$  denotes partial trace on  $\mathcal{H}_C$ , and  $X_{AC}$  denotes an operator  $X \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_C)$ , etc.].

*Proof.* From Lemma 1, it suffices to show

$$\text{Tr}_{\text{HS}} \mathcal{C}_\omega \circ \mathcal{N} \circ \mathcal{J}_\omega = \text{Tr} [R_\omega(\text{id} \otimes \mathcal{N})(S)]. \quad (56)$$

For any linear maps  $\mathcal{E}_1$  and  $\mathcal{E}_2$  on  $\mathcal{L}(\mathcal{H})$  such that  $\mathcal{E}_1$  and  $\mathcal{E}_2^*$  commute, one has

$$(\mathcal{E}_1 \otimes \mathcal{E}_2)^*(S) = (\mathcal{E}_2 \otimes \mathcal{E}_1)(S), \quad (57)$$

where an asterisk denotes the dual linear map. Eq. (57) is seen by (12) because both of  $\text{Tr}[(A \otimes B)(\mathcal{E}_1 \otimes \mathcal{E}_2)^*(S)]$  and  $\text{Tr}[(A \otimes B)(\mathcal{E}_2 \otimes \mathcal{E}_1)(S)]$  are equal to  $\text{Tr}[\mathcal{E}_1(A)\mathcal{E}_2(B)]$ . One has

$$\begin{aligned} \text{Tr}_{\text{HS}} \mathcal{C}_\omega \circ \mathcal{N} \circ \mathcal{J}_\omega &= \text{Tr}_{\text{HS}} \mathcal{J}_\omega \circ \mathcal{C}_\omega \circ \mathcal{N} \\ &= \text{Tr} [S(\text{id} \otimes (\mathcal{J}_\omega \circ \mathcal{C}_\omega \circ \mathcal{N})) (S)] \\ &= \text{Tr} [R_\omega(\text{id} \otimes \mathcal{N})(S)], \end{aligned} \quad (58)$$

where  $R_\omega = ((\mathcal{J}_\omega \circ \mathcal{C}_\omega) \otimes \text{id})(S)$  is the partial transpose on  $\mathcal{H}_B$  of the Choi operator of  $\mathcal{J}_\omega \circ \mathcal{C}_\omega$  and have used the cyclic property of the trace, (15) and (57). Furthermore, one has

$$\begin{aligned} R_\omega &= (\mathcal{J}_\omega \otimes \text{id})[(\mathcal{C}_\omega \otimes \text{id})(S^{TB})]^{TB} \\ &= (\mathcal{J}_\omega \otimes \text{id})(C_\omega^{TB}) \\ &= \text{Tr}_C [(1_A \otimes C_{\omega, CB}^{T_C T_B})(I_{\omega, AC} \otimes \text{id}_B)], \end{aligned} \quad (59)$$

where  $T_B, T_C$  denotes the partial transpose on  $\mathcal{H}_B, \mathcal{H}_C$  and we have used the fact that  $S^{TB}$  is an unnormalized maximally entangled state  $|\Psi\rangle\langle\Psi|$ ,  $|\Psi\rangle = \sum_i |ii\rangle$  and (A.2).  $\square$

By Lemma 4, the problem of finding optimal  $\bar{F}$  and  $\{(\mathcal{J}_\omega, \mathcal{C}_\omega)\}$  is recast in the following form:

$$\begin{aligned} \text{maximize} \quad & f := \sum_{\omega=1}^d \text{Tr} [R_\omega(\text{id} \otimes \mathcal{D}_\varepsilon)(S)] \\ \text{subject to} \quad & C_\omega, I_\omega \geq 0, \quad P = 0, \end{aligned} \quad (60)$$

where  $R_\omega$  is defined in (55) and

$$P := \sum_{\omega} \text{Tr} [(\text{Tr}_A C_\omega - 1)^2] + \text{Tr} [(\text{Tr}_A \sum_{\omega} I_\omega - 1)^2] \quad (61)$$

is the “penalty function.” The conditions  $C_\omega, I_\omega \geq 0$  imply that complete positivity of  $\mathcal{C}_\omega, \mathcal{J}_\omega$  and the condition  $P = 0$  ensures the trace-preserving property of  $\mathcal{C}_\omega$  and  $\sum_{\omega} \mathcal{J}_\omega$  (see the Appendix). When  $\varepsilon = 0, 1$ , the problem can be exactly solved for general  $d = \dim \mathcal{H}$ . For  $\varepsilon = 0$ , the “do nothing” protocol obviously attains the maximum  $\bar{F} = 1$ . For  $\varepsilon = 1$ , it is easy to see that all the protocols fall into “discriminate & reprepare” protocols defined by POVMs. The maximal average fidelity achieved by such protocols is  $\bar{F} = 2/(d+1)$  [23].

One can solve the maximization problem (60) in the following steps.

1. Generate  $2d$  lower triangular matrices  $\{L_{C_\omega}, L_{I_\omega}\}_{1 \leq \omega \leq d}$  so that its nontrivial components are random numbers which obey uniform distribution in the interval  $[-\sqrt{d}, \sqrt{d}]$ . The last is a necessary condition for  $P = 0$ .
2. Set the Choi operators as  $C_\omega = L_{C_\omega} L_{C_\omega}^\dagger$  and  $I_\omega = L_{I_\omega} L_{I_\omega}^\dagger$ . Then  $C_\omega, I_\omega \geq 0$  automatically hold.
3. Apply a numerical maximization method to  $f - \lambda P$ , where  $\lambda$  is a (large) positive number. The penalty term  $-\lambda P$  effectively ensures the condition  $P = 0$ .

We examined Conjecture 1 when  $\dim \mathcal{H} = 3$ . We carried out the numerical scheme above for 5000 initial random points, with  $\lambda = 10^3$  and the maximization method being the simulated annealing. By randomness in the initial data and in the optimization scheme, we expect that the global maximum of  $\bar{F}$  are found. Fig. 3 shows the optimal average fidelity  $\bar{F}$  as a function of the noise strength  $\varepsilon$ . The results suggest that either the “do nothing” or “discriminate & reprepare” protocol is optimal, depending on the strength of the noise. This provides evidence for Conjecture 1.

## VIII. CONCLUSION AND DISCUSSIONS

We have discussed the problem of protecting completely unknown states against given noise by ex-ante and ex-post control scheme. A protocol in the scheme is described mathematically by  $\{(\mathcal{J}_\omega, \mathcal{C}_\omega)\}_{\omega \in \Omega}$  where  $\{\mathcal{J}_\omega\}$  is the CP instrument with the set  $\Omega$  of outcomes applied



before the system goes through the noise and  $\mathcal{C}_\omega$  are the TPCP maps applied after the system suffered the noise. To evaluate the closeness of the input and output states, we have chosen the average fidelity  $\bar{F}$  between the input and output states, which is linear in  $\mathcal{J}_\omega$  and  $\mathcal{C}_\omega$ . We have shown in Theorem 1 that when the scheme involves ex-post control only, one essentially cannot suppress any given noise. In other words, all one can do is to cancel the unitary rotational part of the noise, if one is completely ignorant of the input state. Next, we have considered ex-ante-ex-post control scheme, focusing on the depolarizing noise. We have shown in Theorem 2 that the optimal average fidelity is achieved by protocols which are not “truly quantum,” or can be understood classically. Namely, if the noise is weak, the “do nothing” protocol is optimal, which does nothing to the system literally, and no other protocol can make a larger average fidelity. If the noise is strong, the “discriminate & reprepare” protocol is optimal. In the protocol, one completely measures the system beforehand, discards the resulting state and reconstructs the state estimated from the measurement. The theorems above are for the two-dimensional Hilbert space. Finally, we have proposed Conjecture 1 that Theorem 2 is essentially true for any dimensionality, or more precisely, the optimal average fidelity is achieved either by the “do nothing” or “discriminate & reprepare” protocol. We have found numerical evidence to support the conjecture in the three-dimensional Hilbert space.

A natural question to ask is whether the result similar to Theorem 2 holds or not for noise other than the depolarizing noise. This is not the case at least for the amplitude damping noise, e.g. spontaneous emission; it is shown numerically that there exists a quantum protocol that can do better than the “do nothing” and “discriminate & reprepare” protocols [12]. Our result suggests that noise suppression for completely unknown input states is impossible except by using the bias or anisotropy of the noise itself even when one includes the ex-ante control in the scheme. We note that the depolarizing noise is isotropic in the sense that it commutes with arbitrary unitary operations. Thus our result may be understood as describing the fundamental limitations of quantum mechanics from the viewpoint of noise suppression. It is worth pursuing which class of noise allows nontrivial suppression and identifying the optimal controls therein. In particular, it may be important to examine whether Theorem 2 can be extended to all unitary noise, the noise that preserves the completely mixed state. Our numerical calculations suggest that it is true at least for the dephasing noise.

Let us discuss our results further in the fundamental aspect: irreversibility of quantum processes. While unitary operations describe reversible processes only, TPCP maps include irreversible processes. Then it is natural to ask to what extent a given TPCP map has irreversibility. For an operator on a Hilbert space, the polar decomposition extracts its “irreversible” part uniquely. However, as far as we know, there is no such a simple and canonical

decomposition for TPCP maps, which are operators on Banach spaces. This fact makes it difficult to define the irreversible part of a given TPCP map. In some sense, our work is an attempt to address this problem from the viewpoint of control theory. Using ex-ante and ex-post, we try to cancel an effect of a given TPCP map (noise) and define operationally the irreversible part as what still remains. In Theorem 1, we sought approximate left inverse of a given TPCP map and find it to be a unitary operation. It is not trivial that the approximate left inverse is unique and reversible, which is the conclusion of our theorem. In Theorem 2, using both ex-ante and ex-post control, we found that the irreversible part of the depolarizing noise is itself when the noise is weak.

## ACKNOWLEDGMENTS

T.K acknowledges the support from MEXT-Supported Program for the Strategic Research Foundation at Private Universities Topological Science and from Keio University Creativity Initiative Quantum Community.

## Appendix: Theorems used in the text

In this Appendix, we shall quote some mathematical facts used in the main text.

A square matrix is said doubly stochastic if all components are nonnegative and if the sum over any row is unity and the sum over any column is unity.

**Theorem** (Birkoff-von Neumann [24]). *Any doubly stochastic matrix is a convex combination of permutation matrices.*

There is one-to-one correspondence between CP maps and positive operators on a larger Hilbert space.

**Theorem** (Choi [25]). *Let  $\mathcal{H}_A$  be a  $d$ -dimensional Hilbert space and let  $\mathcal{H}_B$  be a copy of  $\mathcal{H}_A$ . Let  $\{|i\rangle\}_{i \in \{1,2,\dots,d\}}$  be an orthonormal basis for each of  $\mathcal{H}_A$  and  $\mathcal{H}_B$ . Then there is a one-to-one correspondence between a CP map  $\mathcal{E} \in \mathcal{L}(\mathcal{L}(\mathcal{H}_A))$  and a positive operator  $E \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$  such that*

$$\mathcal{E}(\rho) = \text{Tr}_B [(1 \otimes \rho^T)E], \quad (\text{A.1})$$

where  $T$  denotes the transpose with respect to the basis above. The operator  $E$  is called the Choi operator for  $\mathcal{E}$  and is explicitly written as

$$E := (\mathcal{E} \otimes \text{id})(|\Psi\rangle\langle\Psi|) \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B), \quad (\text{A.2})$$

where  $|\Psi\rangle := \sum_i |ii\rangle$  is an unnormalized maximally entangled state. The CP map  $\mathcal{E}$  is trace-preserving if and only if  $\text{Tr}_A E = 1$ .

Let  $\mathcal{H}$  be two-dimensional and  $\mathcal{E}$  be a Hermiticity-preserving linear map on  $\mathcal{L}(\mathcal{H})$ . By a parametrization

$$\mathcal{E}\left(\sum_{\mu=0}^3 x^\mu \sigma_\mu\right) = \sum_{\mu=0}^3 x'^\mu \sigma_\mu, \quad (\text{A.3})$$

$\mathcal{E}$  is expressed as a linear map  $\mathbb{R}^4 \rightarrow \mathbb{R}^4$ ,  $(x^\mu) \mapsto (x'^\mu)$ ,  $\mu = 0, 1, 2, 3$ . If  $\mathcal{E}$  is positive and trace-preserving, then there exist real  $3 \times 3$  rotation matrices  $R, \tilde{R}$  and real numbers  $d_i, t_i$  ( $i = 1, 2, 3$ ) such that

$$|t_i| + |d_i| \leq 1, \quad (\text{A.4})$$

$$\begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} 1 & & & \\ R \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix} & R \begin{pmatrix} d_1 & & \\ & d_2 & \\ & & d_3 \end{pmatrix} & \tilde{R} \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}. \quad (\text{A.5})$$

Ruskai *et al.* gave the concrete parametrization of TPCP maps when  $\mathcal{H}$  is two-dimensional, extending the work by Fujiwara and Algoet [26].

**Theorem** (Ruskai-Szarek-Werner [27], Corollary 2 and

Theorem 4). *Let  $\mathcal{E}$  a positive, trace-preserving linear map specified by (A.3), (A.4) and (A.5).*

(i) *The map  $\mathcal{E}$  is completely positive if and only if all of the following inequalities hold:*

$$(d_1 + d_2)^2 \leq (1 + d_3)^2 - t_3^2 - (t_1^2 + t_2^2) \left( \frac{1 + d_3 \pm t_3}{1 - d_3 \pm t_3} \right), \quad (\text{A.6})$$

$$(d_1 - d_2)^2 \leq (1 - d_3)^2 - t_3^2 - (t_1^2 + t_2^2) \left( \frac{1 - d_3 \pm t_3}{1 + d_3 \pm t_3} \right), \quad (\text{A.7})$$

$$\begin{aligned} & [1 - (d_1^2 + d_2^2 + d_3^2) - (t_1^2 + t_2^2 + t_3^2)]^2 \\ & \geq 4 [d_1^2(t_1^2 + d_2^2) + d_2(t_2^2 + d_3^2) + d_3^2(t_3^2 + d_1^2) - 2d_1d_2d_3]. \end{aligned} \quad (\text{A.8})$$

(ii) *The map  $\mathcal{E}$  is in the closure of the set of extreme points of the space of TPCP maps if and only if there exist  $R, \tilde{R}$  and  $d_1, d_2 \in [-1, 1]$  such that*

$$d_3 = d_1d_2, \quad t_1 = t_2 = 0, \quad t_3^2 = (1 - d_1^2)(1 - d_2^2). \quad (\text{A.9})$$

- 
- [1] W. K. Wootters, W. H. Zurek, *Nature* **299**, 802 (1982); D. Dieks, *Phys. Lett. A* **92**, 271 (1982); H. P. Yuen, *Phys. Lett. A* **113**, 405 (1986).
  - [2] V. Bužek, M. Hillery, *Phys. Rev. A* **54**, 1844 (1996).
  - [3] R. F. Werner, *Phys. Rev. A* **58**, 1827 (1998).
  - [4] C. W. Helstrom, *Inf. Control* **10**, 254 (1967).
  - [5] I. D. Ivanovic, *Phys. Lett. A* **123**, 257 (1987); D. Dieks, *Phys. Lett. A* **126**, 303 (1988); A. Peres, *Phys. Lett. A* **128**, 19 (1988).
  - [6] M. Ozawa, *Ann. Phys.* **311**, 350 (2004).
  - [7] Y. Watanabe, T. Sagawa, and M. Ueda, *Phys. Rev. A* **84**, 042121 (2011).
  - [8] A. M. Brańczyk, P. E. M. F. Mendonça, A. Gilchrist, A. C. Doherty, and S. D. Bartlett, *Phys. Rev. A* **75**, 012329 (2007).
  - [9] P. E. M. F. Mendonça, A. Gilchrist, and A. C. Doherty, *Phys. Rev. A* **78**, 012319 (2008).
  - [10] S. Zhang, X. Zou, C. Li, C. Jin, and G. Guo, e-print arXiv:0811.3254 (2008).
  - [11] A. N. Korotkov and K. Keane, *Phys. Rev. A* **81**, 040103 (2010).
  - [12] C. Q. Wang, B. M. Xu, J. Zou, Z. He, Y. Yan, J. G. Li, and B. Shao, *Phys. Rev. A* **89**, 032303 (2014).
  - [13] P. W. Shor, *Phys. Rev. A* **52**, R2493 (1995).
  - [14] E. Knill and R. Laflamme, *Phys. Rev. A* **55**, 900 (1997).
  - [15] M. Reimpell and R. F. Werner, *Phys. Rev. Lett.* **94** 080501 (2005); N. Yamamoto, S. Hara, and K. Tsumura, *Phys. Rev. A* **71**, 022322 (2005).
  - [16] D. A. Lidar, I. L. Chuang, and K. B. Whaley, *Phys. Rev. Lett.* **81**, 2594 (1998).
  - [17] L. Viola, E. Knill and S. Lloyd, *Phys. Rev. Lett.* **82**, 2417 (1999).
  - [18] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
  - [19] M. Ozawa, *J. Math. Phys.* **25**, 79 (1984).
  - [20] M. Hayashi, *Quantum information: An Introduction* (Springer-Verlag, Berlin, 2006).
  - [21] S. Popescu, A. J. Short, and A. Winter, e-print arXiv:quant-ph/0511225 (2005).
  - [22] S. Massar and S. Popescu, *Phys. Rev. Lett.* **74**, 1259 (1995).
  - [23] G. Chiribella, “On quantum estimation, quantum cloning, and finite quantum de Finetti theorems,” *Theory of Quantum Computation, Communication, and Cryptography*, Lecture Notes in Computer Science, **6519**, 9 (2011).
  - [24] G. Birkhoff, *Univ. Nac. Tucumán Rev. Ser. A* **5**, 147 (1946); J. von Neumann, *Ann. Math. Studies* **28**, 5 (1953).
  - [25] M. D. Choi, *Lin. Alg. Appl.* **10**, 285 (1975).
  - [26] A. Fujiwara and P. Algoet, *Phys. Rev. A* **59**, 3290 (1999).
  - [27] M. B. Ruskai, S. Szarek, and E. Werner, *Lin. Alg. Appl.* **347**, 159 (2002).